MOLECULAR EPIDEMIOLOGY AND GENETIC MARKERS

Prof. Arjumand S. Warsy
Department of Biochemistry
College of Science, King Saud University, Riyadh

Introduction

Recombinant DNA technology has had an enormous impact on all branches of medicine and several epidemiological investigations are now conducted at molecular level. This has led to the emergence of a new field referred to as "Molecular Epidemiology". This is defined as a "Science that focuses on the contribution of potential genetic and environmental risk factors, identified at the molecular and biochemical level, to the etiology, distribution and prevention of disease within families and across populations".

- It differs from traditional genetic epidemiology in that it is equally focused on both environmental and inherited causes of diseases.

- The origin of molecular epidemiology was during the 1980 in cancer research and was described as "an approach in which advanced laboratory methods are used in combination with analytical epidemiology to identify, at the biochemical and molecular level, specific exogenous agents and/or host factors that play a role in human cancer causation".

- More recently, molecular epidemiology has been considered in a broader perspective and is now been used to study numerous infections and chronic diseases with or without genetic basis.

- The most important aim of molecular epidemiology is to identify individuals with a genetic predisposition to a disease, by identify specific disease markers at the DNA level. "DNA markers are actively been sought for all chronic diseases and represent a growing field of DNA diagnostics".

- By incorporation of the recently developed molecular biology technology into epidemiologic research, it has been possible to replace the traditional measures of host susceptibility by more specific genetic markers.

- The field is still young and genetic markers are still not known for several common disorders. On the other hand, these markers exhibit considerable variation in frequency at which they occur in different normal and disease populations.

- Among the various diseases that have been investigated so far at the DNA level in an epidemiological perspective, include insulin dependent diabetes mellitus (IDDM), non-insulin dependent diabetes mellitus (NIDDM), coronary artery disease, hypertension, sickle cell disease, α-thalassaemias, ß-thalassaemias, cancer, hepatitis and several others.

- Molecular Epidemiology involves:

  - Selection of a population for study (normal and disease) i.e. random selection of individuals or families.
  - Extraction of blood for extraction of DNA from the buffy coat.
  - Study of specific marker in the population (normal and disease) using molecular biology techniques.
  - Comparison of the frequency of the occurrance of the marker in the normal population and in the patients with disease, by application of appropriate statistical methods.
  - As a preliminary step it is essential to identify `genetic markers' for various diseases.
  - Some examples of molecular epidemiological studies include:

    (i) Screening for Hb S gene at the DNA level using Mst II:
    This is a very accurate method for diagnosis of presence of $ß^s$ gene. However, since sickle cell haemoglobin in blood can be easily identified on electrophoresis, the need for such molecular epidemiology study has not been felt by most workers. However, this application is of sufficient value when only a small quantity of blood is available (e.g. blood spot on a filter paper) or blood is not available, but other biological material are available e.g. buccal scrapping, hair roots, amniotic fluid, chorionic villus etc., or screening for prenatal diagnosis.

    (ii) Molecular epidemiological studies on ß-thal. patients to identify the mutations causing ß-thal. in different population: In general, 5-8 mutations account for almost 90% of the ß-thal. mutations in each population. Though they are specific for each population and differ from one population to another.

    (iii) Identification of potential risk factors for diabetes mellitus in different populations: Various genetic markers have been investigated for diabetes mellitus (DM). DM is a multifactorial disorder, which develops in genetically susceptible individuals in response to non-genetic environmental factors including life style (i.e. lack of exercise, smoking, excessive eating habits, nature of diet), obesity, viral infections and chemical toxins. Polygenic aetiology has been implicated in its pathogenesis and several studies have been conducted to look for the possible genes involved in its aetiology. The possible genes in the aetiology of DM in the HLA-DR, HLA-DQ, insulin gene and others.

    Of interest are the molecular epidemiological studies in HLA DQ region to study association, if any, with IDDM.

HLA DQ B&A produce ß and α chains. It has been shown that individuals with a codon for any amino acid, other than Asp at codon 57 have a marked resistance to develop IDDM. On the other hand, presence of codon for any other amino acid (non-Asp 57) provides a significant susceptibility.

Similarly presence of an Arginine codon at position 52 of HLA DQ A confers susceptibility, while any other amino acid at this positions confers resistance.

- Thus Asp 57 and Arg 52 determination in the HLA DQ gene have been the subject of extensive studies in different populations. In Saudi IDDM patients, similar studies have shown that frequency of Asp 57 is higher in non-diabetis compared to IDDM while the non-Asp 57 occur at a higher frequency in IDDM cases.

- On the other hand, Arg 52 occurs at a higher frequency in IDDM patients compared to their non diabetic control.

- The hypervariable region on 5' end of insulin gene on chromosome 6 has also been shown to have some association with IDDM development. However, the data from different populations show significant differences.

- Other markers for IDDM include HLA DR 3 and HLA DR4 which increase the risk of IDDM, while HLA DR-2 provides resistance against IDDM development.

- Several markers have been investigated for NIDDM and some have proved to be of value in a few populations. These include glucose transport gene, insulin receptor genes and apolipoproten genes.

- Several studies are in progress to search for preclinical markers for genetic disorder's particularly those with - multifactorial aetiology.